



Caractérisation d'experts dans les plate-formes de crowdsourcing Characterization of experts in crowdsourcing platforms

Amal Ben Rjab, Mouloud Kharoune, Zoltan Miklos, Arnaud Martin,
Boutheina Ben Yaghlane

► To cite this version:

Amal Ben Rjab, Mouloud Kharoune, Zoltan Miklos, Arnaud Martin, Boutheina Ben Yaghlane. Caractérisation d'experts dans les plate-formes de crowdsourcing Characterization of experts in crowdsourcing platforms. 24ème Conférence sur la Logique Floue et ses Applications, Nov 2015, Poitiers, France. hal-01263635

HAL Id: hal-01263635

<https://inria.hal.science/hal-01263635>

Submitted on 27 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Caractérisation d'experts dans les plate-formes de crowdsourcing

Characterization of experts in crowdsourcing platforms

A. Ben Rjab¹

M. Kharoune²

Z. Miklos²

A. Martin²

B. Ben Yaghlane³

¹ Université de La Manouba / École supérieure de commerce de Tunis

² Université de Rennes 1/ IRISA

³ Université de Carthage / Institut des hautes études commerciales de Carthage

benrjabamal.ihec@gmail.com

Mouloud.Kharoune@univ-rennes1.fr

zoltan.miklos@irisa.fr

Arnaud.Martin@univ-rennes1.fr

boutheina_yaghlane@yahoo.fr

Résumé :

Le "Crowdsourcing" représente aujourd'hui une méthode de travail très populaire qui a une grande importance. Cette méthode consiste à recourir aux ressources de la foule pour accomplir des tâches, pour collecter des informations et des connaissances sous forme d'une sous-traitance vers un grand nombre de personnes. La qualité et le niveau d'expertise de ces derniers influencent largement l'efficacité des plates-formes de crowdsourcing. Ceci pose un nouveau défi de contrôle de qualité des participants et d'identifier les experts. Plusieurs méthodes ont été proposées dans ce contexte, mais les problèmes d'imperfection des données souvent peu prises en compte influencent tous les résultats observés. Pour ce faire et afin de mieux identifier les experts nous proposons une nouvelle méthode de caractérisation des experts fondée sur la théorie des fonctions de croyance. Nous validons nos techniques par des expérimentations sur des données synthétiques.

Mots-clés :

Crowdsourcing, Expert, Théorie des fonctions de croyance, Exactitude.

Abstract:

Crowdsourcing represents a very popular working method which acquires a great importance. This method consists to recourse to the crowd to accomplish tasks, to collect pieces of information and knowledge through outsourcing the tasks to a large number of people. The quality and the level of expertise of participants strongly influence the efficacy of the crowdsourcing platforms. This raises new challenges to control the quality of participants and identifying experts. In this context, several methods have been proposed, but the problem of imperfection data is often not taken into account and this might largely influence all the observed results. We propose novel methods for identifying the experts, based on the theory of belief functions. We validate our techniques through experimentation on synthetic datasets.

1 Introduction

Au cours de ces dernières années et avec l'émergence du web, le crowdsourcing a connu une forte progression. Il est considéré aujourd'hui comme l'une des meilleures approches de travail [1], en particulier pour les tâches qui sont faciles à effectuer pour des utilisateurs humains mais difficiles pour un ordinateur. Cette notion popularisée par Howe¹ en 2006 [2], conquiert aujourd'hui le monde par ces techniques de partage, de collection d'informations et de résolution de problèmes et en fait une des méthodes de sollicitation les plus puissantes [3]. Le crowdsourcing ou l'acte d'externalisation des tâches permet de recourir aux ressources de la foule pour résoudre des problèmes, pour collecter des informations et pour distribuer des connaissances ou des savoirs à une grande échelle [4]. Jeff Howe a considéré ce concept en 2006 dans son article [2] comme l'acte d'externalisation à un grand nombre de personnes sous forme d'un appel ouvert et en 2008 [5] il l'a défini lui-même comme "*the act of taking a job traditionally performed by a designated agent and outsourcing it to an undefi-*

1. Jeff Howe est un éditeur contribuant dans Wired-magazine, qui le premier, introduit le terme de crowdsourcing et l'étudia *via* des recherches scientifiques.
<http://crowdsourcing.typepad.com>

ned, generally large group of people in the form of an open call". Cependant, malgré le succès de ces plates-formes, les questions relatives à la qualité des réponses obtenues par la foule et le contrôle de cette qualité pour permettre d'identifier les experts reste encore un défi. Plusieurs méthodes [1, 4, 6, 7, 8] ont été proposées pour résoudre ce problème d'identification des experts. Cependant ces approches manipulent des données très imparfaites, sans pour autant chercher à modéliser ces imperfections. Il est pourtant important de bien modéliser ces imperfections issues des différentes sources afin de garantir une identification la plus précise et fiable. Dans ce contexte, nous proposons une nouvelle méthode d'identification des experts fondée sur la théorie des fonctions de croyance permettant de bien modéliser les imperfections des réponses des participants et d'estimer le degré d'expertise.

Une analyse de l'état de l'art sur l'identification des experts est présentée dans la section 2. Les notions d'expert et d'expertise sont détaillées dans la section 2.1. La section 3 présente un rappel sur la théorie des fonctions de croyance, avant de présenter notre approche pour définir un degré d'exactitude dans la section 4 et de l'illustrer par des expérimentations dans la dernière section.

2 Identification des experts dans le crowdsourcing

Dans un environnement difficile et compétitif comme celui de crowdsourcing où l'innovation et l'expertise font la différence, il est important d'être attentif aux attentes des clients. Évaluer la qualité des participants et identifier les experts parmi eux dans les plates-formes de crowdsourcing reste encore un problème délicat.

2.1 Notion d'expert dans le contexte crowdsourcing

L'identification de l'expertise d'un participant ne peut pas être considérée comme un nouveau sujet, la nécessité de reconnaître les personnes à fort capacité intellectuel, technique et scientifique a été abordée dans tous les domaines et dans toutes les plate-formes.

Un expert dans les plate-formes de crowdsourcing n'est pas simplement celui qui sait, mais plutôt celui qui a acquis un ensemble de connaissances, de compétences, de savoirs et de créativité. De fait, c'est celui qui fournit un grand nombre de réponses correctes [9], précises, complètes et fiables pour un sujet spécifique. L'expert doit être capable de fournir des résultats efficaces et fiables [1] et d'exprimer des jugements pour tous les demandeurs, c'est pourquoi plusieurs auteurs l'ont défini comme la personne qui a des responsabilités spécifiques [10]. Les experts sont identifiés dans cet environnement par :

- La précision et l'exactitude de ces réponses
- Leur capacité de détecter et de reconnaître les tâches *a priori*
- Leur capacité d'extraire les données et les solutions avec un minimum d'effort cognitif
- Leur capacité de contrôle et d'évaluation
- Leur connaissance, leur motivation, leur réflexion et leur capacité d'apprentissage

Généralement on peut identifier l'expert par le rôle qu'il joue, par les natures des tâches qu'il réalise, par le temps de résolution du problème et par les pertinences des solutions proposées qui reflètent le niveau de connaissances et d'expertise du participant.

2.2 Méthodes d'identification d'experts

Plusieurs auteurs ont exploré ce domaine et plusieurs méthodes ont été proposées dans ce cadre. Récemment Dawid et Skene ont proposé dans [11] un modèle de classification latente pour estimer la qualité des participants, *via* un taux d'erreurs fondé sur le maximum de vraisemblance. Ipeirotis *et al.* [1] ont modifié l'algorithme (EM) proposé par Dawid séparant entre les deux taux de biais et d'erreur, afin de générer un score qui représente la qualité de chaque participant. Un autre algorithme fondé sur les graphes : SPEAR (spamming resistant Expertise Analysis and Ranking) est proposé par Noll *et al.* en 2009 [7]. Guittard et Schenk dans [4] ont proposé une classification des participants avec les technologies de data mining et la méthode du vote pour identifier les experts. Khattak et Aouissi ont proposé dans [8] une évaluation des participants selon l'ensemble des étiquettes fournies, l'idée est d'offrir un cadre appelé "Expert Label Injected Crowd Estimation" (ELICE) consistant à collecter toutes les étiquettes fournies par les participants et les comparer à celles qui est la meilleure nommée *label expert*. Raykar *et al.* ont proposé dans [6] d'utiliser un estimateur du maximum de vraisemblance alors que Kazai *et al.* dans [12] se sont fondés sur des observations comportementales pour définir une typologie des travailleurs en *Spammer, Sloppy, Incompetent, Competent, Diligent*.

Et puisque le problème d'identification des experts n'est pas seulement relatif aux plateformes de crowdsourcing, d'autres méthodes ont été proposées dans des différents domaines, dans les communautés en ligne de Java Forum; Zhang *et al.* dans [13] ont proposé une analyse d'expertise fondée sur le calcul d'un score à l'aide de l'algorithme de PageRank (proposé initialement par Page en 1998 pour organiser les pages web). Ils ont l'amélioré par un calcul de score nommé "z-score" qui consiste à combiner le nombre de réponses (a) et le nombre de questions (q) tels que :

$Z = \frac{a - n/2}{\sqrt{n/2}} = \frac{a - q}{\sqrt{a + q}}$. Dans la plate-forme Yahoo-Answers, Bouguessa *et al.* [14] ont proposé un modèle pour identifier les utilisateurs ayant autorité fondé sur le nombre de bonnes réponses. Dans le domaine des réseaux sociaux, pour choisir les experts, Bozzon *et al.* [15] ont proposé un classement des membres des groupes sociaux selon leur niveau de connaissance qui est déduit par une analyse de leurs profils et des traces de leurs activités sociales.

Nous constatons ainsi que d'une part la notion d'expert n'est pas perçue de la même façon selon les approches et d'autre part que ces approches manipulent des données très imparfaites pour détecter des niveaux d'expertise, sans pour autant chercher à modéliser ces imperfections. Pour aborder ce dernier point, nous proposons d'employer la théorie des fonctions de croyance, dont nous rappelons les bases dans la section suivante.

3 Théorie des fonctions de croyance

La théorie des fonctions de croyance issue des travaux de Dempster [16] et Shafer [17] permet une bonne modélisation des données imprécises et incertaines et offre des outils puissants pour modéliser, fusionner et gérer les informations issues de plusieurs sources.

Cette théorie est fondée sur la manipulation des fonctions de masse qui sont définies sur l'ensemble de toutes les disjonctions du cadre de discernement. Le cadre de discernement $\Omega = \{w_1, \dots, w_n\}$ représente l'ensemble de toutes les hypothèses exclusives et exhaustives. Ainsi, une fonction de masse m^Ω est définie sur l'espace puissance $2^\Omega = \{A | A \subseteq \Omega\} = \{\emptyset, \omega_1, \omega_2, \omega_3, \dots, \omega_n, w_1 \cup \omega_2, \Omega\}$ et à valeurs dans $[0,1]$, représente le degré de croyance élémentaire sur ces éléments, et vérifie :

$$\sum_{A \subseteq \Omega} m^\Omega(A) = 1 \quad (1)$$

À partir de la fonction de masse, d'autres fonctions de croyance peuvent être définies afin de modéliser incertitude et imprécision, telles que les fonctions de plausibilité et crédibilité.

Différentes règles de combinaison ont été proposées pour fusionner les informations issues de différentes sources. Initialement Dempster a introduit la règle de combinaison conjonctive normalisée pour chercher le consensus entre plusieurs sources fiables et cognitivement indépendantes. Elle est donnée par :

$$m_D(X) = \frac{1}{1-k} \sum_{Y_1 \cap Y_2 = X} m_1(Y_1)m_2(Y_2) \quad (2)$$

où $k = \sum_{Y_1 \cap Y_2 = \emptyset} m_1(Y_1)m_2(Y_2)$ est l'inconsistance de la combinaison également nommé conflit global. Cependant cette règle a l'inconvénient de la convergence rapide de la valeur k à 1 lorsque le nombre de sources combiné est grand. Dans ce papier la moyenne des fonctions de masse sera donc employée.

Afin de comparer les fonctions de masse entre elles, plusieurs distances on été proposées. Nous retenons ici celle proposée par Jousselme et al. [18] qui permet de tenir compte de la cardinalité différentes de l'espace 2^Ω . Elle est donnée par :

$$d_J^2(m_1, m_2) = \frac{1}{2}(m_1 - m_2)^T \underline{\underline{D}}(m_1 - m_2) \quad (3)$$

où $\underline{\underline{D}}$ est la matrice définie par Jaccard :

$$D(X, Y) = \begin{cases} 1 & \text{if } X = Y = \emptyset \\ \frac{|X \cap Y|}{|X \cup Y|} & , \forall x, y \in 2^\Omega \end{cases}$$

4 Degré d'exactitude des réponses

Comme nous l'avons vu dans la section 2.1, la notion d'expert peut être évaluée selon plusieurs facteurs. Dans la méthode que nous proposons, nous nous limitons à mesurer l'exactitude des réponses des participants.

Afin de modéliser les imperfections des réponses des participants, nous supposons que

chaque réponse représentée par une fonction de masse. Soit U_j un participant caractérisé par ses réponses r_{U_j} qui sont chacune modélisée par une fonction de masse $m_{U_j}^{\Omega_k}$. Chaque réponse est relative à une question Q_k ($k = \{1, \dots, K\}$) qui est défini dans un cadre de discernement spécifique Ω_k avec $\Omega_k = \{\omega_1^{Q_k}, \dots, \omega_{n_k}^{Q_k}\}$. Nous obtenons ainsi une matrice des s participants par K questions donnée par :

$$\begin{matrix} & \begin{matrix} Q_1 & \dots & Q_k & \dots & Q_K \end{matrix} \\ \begin{matrix} U_1 \\ \vdots \\ U_j \\ \vdots \\ U_s \end{matrix} & \begin{bmatrix} m_{U_1}^{\Omega_1} & \dots & m_{U_1}^{\Omega_k} & \dots & m_{U_1}^{\Omega_K} \\ \vdots & & \vdots & & \vdots \\ m_{U_j}^{\Omega_1} & \dots & m_{U_j}^{\Omega_k} & \dots & m_{U_j}^{\Omega_K} \\ \vdots & & \vdots & & \vdots \\ m_{U_s}^{\Omega_1} & \dots & m_{U_s}^{\Omega_k} & \dots & m_{U_s}^{\Omega_K} \end{bmatrix} \end{matrix} \quad (4)$$

Afin de mesurer un degré d'exactitude pour chaque participant, l'idée est de calculer la distance entre la fonction masse $m_{U_j}^{\Omega_k}$ de la réponse proposée par le participant U_j pour la $k^{\text{ème}}$ question Q_k , avec une représentation des réponses des autres participants $m_{U_{\varepsilon_{s-1}}}^{\Omega_k}$ pour cette même question. Cette représentation des réponses $m_{U_{\varepsilon_{s-1}}}^{\Omega_k}$ est donnée ici par la moyenne des fonctions de masse données par les $s-1$ autres participants (sauf le participant U_j sélectionné) :

$$m_{U_{\varepsilon_{s-1}}}^{\Omega_k}(X) = \frac{1}{s-1} \sum_{\substack{i=1 \\ i \neq j}}^{s-1} m_{U_i}^{\Omega_k}(X) \quad (5)$$

L'hypothèse sous-jacente à cette modélisation est le fait que la majorité des participants a raison. La moyenne permet de représenter le centre des fonctions de masse même si elles sont très différentes. Les règles du type conjonctif entraînent rapidement une masse sur l'ensemble vide en fonction du nombre de fonctions de masse combinées. Ainsi pour chaque participant U_j et chaque question Q_k , nous obtenons la distance $d_{U_j}^{\Omega_k}$ donnée par la distance de Jousselme de l'équation (3) : $d_J(m_{U_j}^{\Omega_k}, m_{U_{\varepsilon_{s-1}}}^{\Omega_k})$

Une autre distance que celle de Joussette pourrait être employée. À partir de cette distance, nous mesurons ainsi l'inexactitude des réponses et nous définissons un degré d'exactitude des participants par :

$$IE(U_j) = 1 - \frac{1}{r(U_j)} \sum_{k=1}^K d_{U_j}^{\Omega_k} \quad (6)$$

Sachant que r_{U_j} représente ici le nombre des réponses proposé par le participant U_j pour une question donnée (k), avec $U_j \neq K$ et $d_{U_j}^{\Omega_k}$ représente la distance calculée pour la même question entre $m_{U_j}^{\Omega_k}$ et $m_{U_{\varepsilon_s-1}}^{\Omega_k}$. Afin de déterminer les experts parmi tous les participants, nous appliquons un k -moyenne avec $k=2$. L'utilisation de k -moyenne permet de confirmer les groupes. Avec cette technique l'ensemble des experts sera donné par la classe ayant la moyenne des degrés d'exactitude la plus élevée. Sous l'hypothèse que la majorité a raison, nous avons choisi $k = 2$. Avec cette hypothèse tous les participants qui donnent des réponses exactes doivent être considérés comme experts et les restes comme des ignorants.

5 Expérimentation

Pour vérifier l'efficacité de notre méthode, nous avons appliqué plusieurs tests à partir de données générées. Les fonctions de masse sont directement générées en considérant trois types de participants :

- Les experts : les fonctions de masse sont à support simple, pour élément focal un singleton attendu avec une valeur entre 0.75 et 1.
- Les futurs experts : les fonctions de masse ont deux éléments focaux avec un élément qui contient le singleton attendu et l'ignorance.
- Les ignorants : les fonctions de masse sont non dogmatiques et tirées aléatoirement.

Pour chaque question la taille du cadre de discernement Ω_k est choisie aléatoirement entre 2 et 6. À l'étape d'initialisation notre algorithme permet de préciser le nombre des participants, le nombre des questions ainsi le pourcentage des experts à identifier et des futurs experts. Il permet aussi de préciser le mode des réponses (donne la possibilité d'avoir ou pas des réponses vides). L'algorithme 1 reprend les étapes du calcul du degré d'exactitude qui reflète le niveau d'expertise des participants.

Dans les différentes expérimentations présentées ci-dessous, nous allons considérer 100 participants et 100 questions. Nous réalisons la moyenne sur les taux de bonnes classification après 10 répétitions de l'expérience afin de conforter les résultats.

Dans une première expérience, nous faisons varier le pourcentage d'experts sans générer de futurs experts (0%). Les résultats sont illustrés sur la figure 1.

Nous pouvons constater qu'au dessus de 43 % d'experts nous obtenons plus de 50 % de taux de bonne classification.

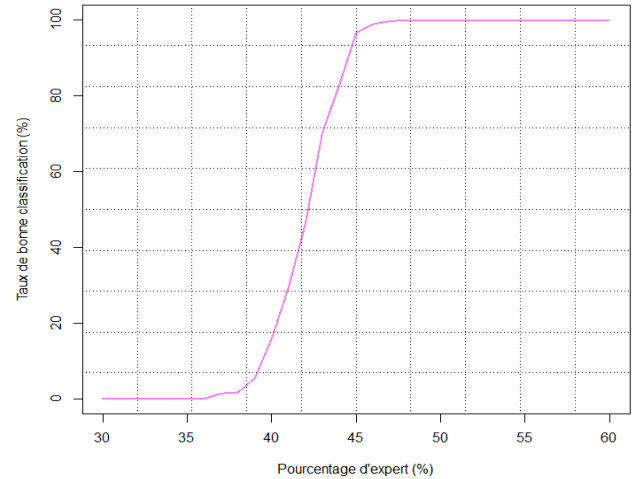


Figure 1 – Représentation de la variation du taux de bonne classification par rapport à la variation des pourcentages des experts

L'hypothèse que la majorité des participants donne la réponse correcte est essentielle. Sous cette hypothèse l'approche proposée permet

Algorithm 1: DEGRÉ D'EXACTITUDE

Input: K : le nombre de questions s : le nombre de participants E : le nombre d'experts F : le nombre de futurs experts**Output:** IE : le degré d'exactitude

```
1  $AllRes \leftarrow$  list des réponses des participants
2  $AllCombine \leftarrow$  list des masses combinées
3  $AllDist \leftarrow$  matricededistance
4 PHASE 1 : Génération des masses
5 for  $k \in [1 : K]$  do
6   for  $u \in [1 : E]$  do
7      $AllRes[k][u] \leftarrow$  réponse d'un expert
8   for  $u \in [(E + 1) : (E + F)]$  do
9      $AllRes[k][u] \leftarrow$  réponse d'un futur expert
10  for  $u \in [(E + F + 1) : s]$  do
11     $AllRes[k][u] \leftarrow$  réponse aléatoire
12 return  $AllRes$ 
13 PHASE 2 : Combinaison des masses
14 for  $k \in [1 : K]$  do
15   for  $u \in [1 : s]$  do
16      $AllCombine[k, u] \leftarrow$ 
17        $meanComb(AllRes[k][u], AllRes[k][s - u + 1])$ 
18 return  $AllCombine$ 
19 PHASE 3 : Calcul des distances
20 for  $k \in [1 : K]$  do
21   for  $u \in [1 : s]$  do
22      $AllDist[k, u] \leftarrow JDistance($ 
23        $AllRes[k][u],$ 
24        $AllCombine[k, u])$ 
25 return  $AllDist$ 
26 PHASE 4 : Calcul du degré : moyenne des distances
27 for  $u \in [1 : P]$  do
28    $IE[u] \leftarrow mean(AllDist[1 : K, u])$ 
29 return IE
```

de bien caractériser les experts des ignorants. Lorsque la majorité des réponses est complètement aléatoire (donnée par les ignorants), ceux-ci sont considérés comme les ex-

perts. La cohérence des réponses des ignorants n'est pas considérée ici.

Dans la deuxième expérimentation nous faisons varier le pourcentage d'experts et de futurs experts de façon identique de 10 % jusqu'à 50 %, afin d'identifier réellement les capacités de discrimination de notre approche.

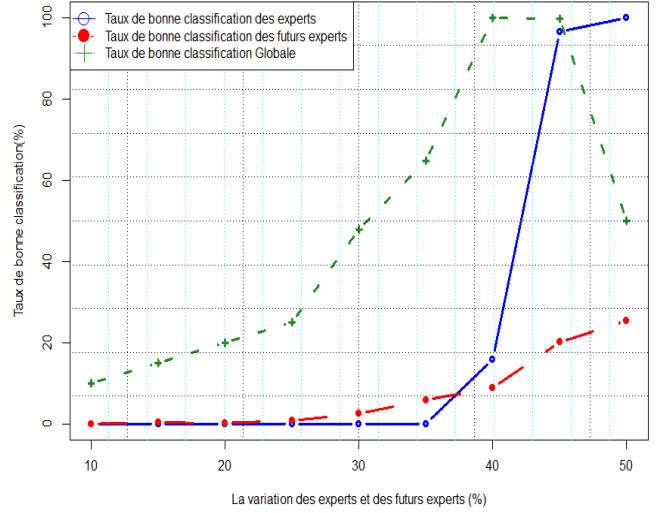


Figure 2 – Représentation de la variation du taux de bonne classification par rapport à la variation des experts et des futurs experts

La figure 2 permet de représenter simultanément les taux de bonnes classification des experts, des futurs experts et le taux global.

Nous constatons que les futurs experts sont mieux discriminés sur les faibles pourcentages jusqu'à 40 %. A 40 % les experts commencent à être tous identifiés correctement. Bien sûr à 50 % pour les experts et futurs experts, il n'y a plus d'ignorants. Notre approche cherchant à caractériser deux classes, les futurs experts sont alors considérés comme des ignorants.

Dans une troisième expérimentation, nous faisons varier le nombre des questions avec 35 % experts et 35 % de futurs experts pour les 10 itérations, afin de vérifier la stabilité des taux de bonne classification. Avec ces pourcentages, nous avons 64.8 % de bonne classification. Pour mesurer cette stabilité, nous calculons un taux de perturbation donné par l'écart type entre

les différents taux de bonne classification sur les 10 itérations. Les résultats sont illustrés par la figure 3.

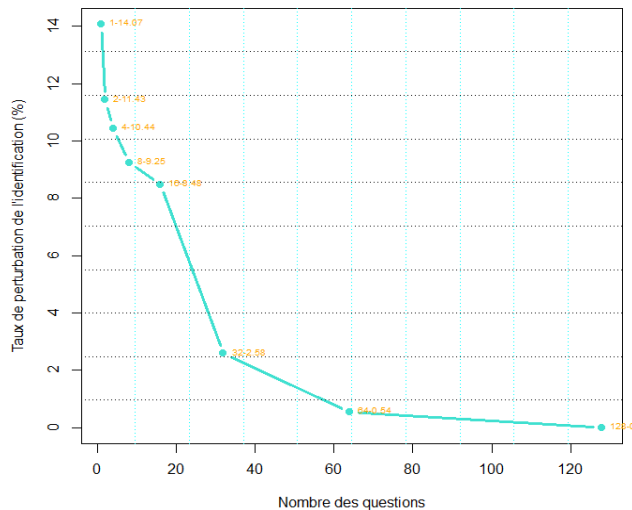


Figure 3 – Représentation de la variation du taux de perturbation de l'identification par rapport à la variation du nombre de questions

Cette figure montre ainsi qu'il est nécessaire d'avoir un certain nombre de questions pour espérer identifier les experts. Ainsi pour 100 questions, les estimations des taux de classification obtenues dans les expérimentations précédentes (Figure 1 et 2) peuvent être considérées comme fiables.

Ces résultats confirment donc l'intérêt d'employer la théorie des fonctions de croyance pour la détection des experts pour des problèmes de crowdsourcing.

6 Conclusion et perspectives

La qualité des participants dans la foule est un facteur clé de succès du crowdsourcing. Identifier les experts parmi tous les participants reste un défi pour ces plate-formes. Plusieurs méthodes ont été proposées pour identifier et caractériser les experts, mais elles ne modélisent pas l'imperfection des réponses issus des différents utilisateurs. Dans cet article

nous avons proposée une méthode d'identification des experts fondée sur un degré d'exactitude des réponses des participants modélisées dans le cadre de la théorie des fonctions de croyance. Modéliser les réponses des participants à l'aide des fonctions de croyance offre la possibilité d'une meilleure maîtrise dans la caractérisation des experts.

La méthode proposée nous permet de bien caractériser trois types de comportement des participants : *Experts*, *futurs experts* et *ignorants* par le calcul d'un degré d'exactitude qui reflète leur niveau d'expertise. L'hypothèse principale de notre approche est le fait que pour une question donnée, la majorité des participants donnent la bonne réponse. Le degré d'exactitude défini permet ainsi de distinguer les experts des autres participants à partir de 40 % experts.

Nous avons, de part la génération des fonctions de masse, différencié les experts (qui donnent une réponse exacte et précise) des futurs experts (qui donnent des réponses exactes mais imprécises). L'étude du pourcentage de réponses exactes reste à faire. Afin de distinguer correctement les réponses précises des réponses imprécises, qui doivent être considérées comme moins expertes, nous envisageons d'ajouter un degré de précision. La précision est une caractéristique importante de la notion d'expert. Il sera donc intéressant de combiner les degrés d'exactitude et de précision des participants afin de caractériser les experts sur des cas plus réalistes que ceux investigués dans ce papier.

Références

- [1] Panagiotis G Ipeirotis, Foster Provost, and Jing Wang. Quality management on amazon mechanical turk. In *Proceedings of the ACM SIGKDD workshop on human computation*, pages 64–67. ACM, 2010.
- [2] Jeff Howe. The rise of crowdsourcing. *Wired magazine*, 14(6) :1–4, 2006.

- [3] Aniket Kittur, Jeffrey V Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1301–1318. ACM, 2013.
- [4] Claude Guittard and Eric Schenk. Le crowdsourcing : typologie et enjeux d’une externalisation vers la foule. *Document de travail du bureau d’économie théorique et appliquée*, 2, 2011.
- [5] Jeff Howe. *Crowdsourcing : How the power of the crowd is driving the future of business*. Random House, 2008.
- [6] Vikas C Raykar, Shipeng Yu, Linda H Zhao, Gerardo Hermosillo Valadez, Charles Florin, Luca Bogoni, and Linda Moy. Learning from crowds. *The Journal of Machine Learning Research*, 11 :1297–1322, 2010.
- [7] Michael G Noll, Ching-man Au Yeung, Nicholas Gibbins, Christoph Meinel, and Nigel Shadbolt. Telling experts from spammers : expertise ranking in folksonomies. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 612–619. ACM, 2009.
- [8] Faiza Khan Khattak and Ansaf Salleb-Aouissi. Quality control of crowd labeling through expert evaluation. In *Proceedings of the NIPS 2nd Workshop on Computational Social Science and the Wisdom of Crowds*, 2011.
- [9] Aditya Pal and Joseph A Konstan. Expert identification in community question answering : exploring question selection bias. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 1505–1508. ACM, 2010.
- [10] Sébastien Dubois, Najoua Mohib, David Oget, Eric Schenk, and Michel Sonntag. Connaissances et reconnaissance de l’expert. *Les cahiers de l’INSA de Strasbourg*, (1) :89–108, 2005.
- [11] Alexander Philip Dawid and Allan M Skene. Maximum likelihood estimation of observer error-rates using the em algorithm. *Applied statistics*, pages 20–28, 1979.
- [12] Gabriella Kazai, Jaap Kamps, and Natasa Milic-Frayling. Worker types and personality traits in crowdsourcing relevance labels. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pages 1941–1944. ACM, 2011.
- [13] Jun Zhang, Mark S Ackerman, and Lada Adamic. Expertise networks in online communities : structure and algorithms. In *Proceedings of the 16th international conference on World Wide Web*, pages 221–230. ACM, 2007.
- [14] Mohamed Bouguessa, Benoît Dumoulin, and Shengrui Wang. Identifying authoritative actors in question-answering forums : the case of yahoo! answers. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 866–874. ACM, 2008.
- [15] Alessandro Bozzon, Marco Brambilla, Stefano Ceri, Matteo Silvestri, and Giuliano Vesci. Choosing the right crowd : expert finding in social networks. In *Proceedings of the 16th International Conference on Extending Database Technology*, pages 637–648. ACM, 2013.
- [16] Arthur P Dempster. Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, pages 325–339, 1967.
- [17] Glenn Shafer. *A mathematical theory of evidence*, volume 1. Princeton university press Princeton, 1976.
- [18] Anne-Laure Jousselme, Dominic Grenier, and Éloi Bossé. A new distance between two bodies of evidence. *Information fusion*, 2(2) :91–101, 2001.

